TRABAJOS DE ESTADISTICA Y
DE INVESTIGACION OPERATIVA
Vol. 34, Núm. 3, 1983, pp. 79 a 94

METODOS DE OBTENCION DE LA INFORMACION ESPERADA GLOBAL

Ernesto Veres Ferrer
Instituto Nacional de Estadística
Teruel

RESUMEN

En este trabajo se acomete una generalización de la definición de Shannon - Lindley para la información esperada proporcionada por un experimento que presupone la existencia de estratificación en el espacio muestral. Ante la evidente dificultad de cálculo de la información esperada en la situación planteada –dificultad que se deriva de la existencia de un vector como parámetro de interés y de un resultado muestral que es un conjunto de muestras obtenidas de poblaciones distintas– en este artículo se presentan dos procedimientos que reducen la situación planteada por la existencia de un experimento asociado a cierto diseño de estratificación a la situación de varias dimensiones de la ya estudiada por Lindley (1956). Finalmente, la coherencia de ambos procedimientos queda manifestada en su aplicación al modelo normal.

SUMMARY

We shall develop in this paper a generalitation of the Shannon - Lindley definition for the expected information provided by an experiment with a stratified sampling space. Due to the expected information calculus difficulty –because to exist a vector parameter or vector quantity of interest and a sampling result who it is many samples, extracted from different poblations– whe shall present two methodes to reduce the situation, because to exist an associated experiment to certain stratified sampling by Lindley (1956) studied many dimensional situation. Endly, the two methods coherence is manifested in their normal model application.

(*) Recibido, Junio, 1981

Palabras y frases clave:

Diseño de experimentos; Estratificación; Inferencia Estadística; Información esperada global; Métodos Bayesianos; Modelo normal.

Números de clasificación de A.M.S. (1970)

- Primarios: 62B10; 62F15.
- Secundarios: 62A15; 62B15; 62K99.

1. INTRODUCCION. INFORMACION ESPERADA GLOBAL

Consideremos una situación de Inferencia Estadística. Así pues, supondremos primeramente que el objetivo pretendido por cierta investigación es el de mejorar nuestro conocimiento sobre el valor de cierta cantidad o parámetro de interés. Para ello, el científico interesado en realizar inferencia diseñará un experimento de forma que la distribución de sus resultados dependerá de la cantidad de interés de una manera conocida, proporcionando la observación de tales resultados cierta cantidad de información.

A fin de describir un experimento asociado a un diseño de estratificación, el modelo matemático básico a considerar contiene un espacio muestral X en el que existe una partición en L subconjuntos o estratos X_i $(i = 1, 2, ..., L)$. Cada uno de éstos está dotado de una apropiada σ-álgebra sobre la que se definen sendas familias de medidas de probabilidad. Las densidades de cada una de ellas se individualizan mediante un parámetro θ_i que toma valores en su respectivo espacio paramétrico Θ_i.

Para cada estrato se supone definido el experimento $E_i = \{X_i, \Theta_i, p(x_i/\theta_i)\}$ que consiste en una observación de la variable aleatoria $x_i \in X_i$ que se distribuye, para algún $\vartheta_i \in \Theta_i$ dado, de acuerdo con la densidad $p(x_i/\theta_i)$. En estas condiciones definiremos el experimento compuesto $E^s(n)$ como aquél que está formado por los L experimentos componentes independientes $E(n_i)$ $(i = 1, 2, ..., L)$, y en donde $E(n_i)$ representa al experimento que consiste en n_i repeticiones independientes (dado θ_i) del experimento E_i. Con $n' = (n_1, n_2, ..., n_L)$ denotamos al vector tamaño muestral cuyas componentes son los respectivos tamaños de las muestras extraídas en sus estratos respectivos.

La densidad de probabilidad que describe el comportamiento de la muestra global z resultado de $E^s(n)$ - muestra que no es sino el conjunto de submuestras parciales $z = (z_1, z_2, ..., z_L)$ extraídas de sendos estratos y que resultan, a su vez, de los respectivos experimentos independientes $E(n_i)$ - viene dada a través del doble productorio

$$p(z/\underset{\sim}{\theta}) = \prod_{i=1}^{L} p(z_i/\theta_i) = \prod_{i=1}^{L} \prod_{j=1}^{n_i} p(x_{ij}/\theta_i)$$

en el que con $\underset{\sim}{\theta}$ denotamos al vector cantidad de interés o vector paramétrico cuyas componentes son las respectivas cantidades de interés θ_i, esto es, $\underset{\sim}{\theta}' = (\theta_1, \theta_2, ..., \theta_L)$ y que toma valores en el conjunto $\underset{\sim}{\Theta} = \Theta_1 \, x \ldots x \, \Theta_L$.

A fin de terminar de describir los elementos del problema, el argumento Bayesiano extiende este modelo básico al suponer que $\underset{\sim}{\Theta}$ es soporte de una apropiada σ - álgebra, de forma que las opiniones iniciales del investigador antes de realizar $E^s(\underset{\sim}{n})$ se describen a través de la densidad de probabilidad inicial $p(\underset{\sim}{\theta})$.

Siendo $p(\underset{\sim}{\theta})$ una densidad inicial, la densidad posterior que describe las opiniones a posteriori del investigador viene dada a través del Teorema de Bayes:

$$p(\underset{\sim}{\theta}/z) = p(z/\underset{\sim}{\theta}) \cdot p(\underset{\sim}{\theta})/p(z)$$

donde $p(z)$ es la correspondiente densidad predictiva. Supondremos en lo que sigue y salvo indicación en contra que cuantas integrales aparezcan existen y están extendidas a sus respectivos dominios de definición completos.

Con el conjunto de elementos anterior podemos definir a la información esperada global como una medida de la información que sobre $\underset{\sim}{\theta}$ proporciona $E^s(\underset{\sim}{n})$ y que supone la generalización a este experimento del correspondiente concepto definido por Lindley (1956), siguiendo el trabajo de Shannon (1948).

DEFINICION. La información esperada global sobre $\underset{\sim}{\theta}$ proporcionada por el experimento $E^s(\underset{\sim}{n})$, cuando la densidad inicial es $p(\underset{\sim}{\theta})$, viene dada a través de la doble integral

$$I^{\theta}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} = \iint p(z) \cdot p(\underset{\sim}{\theta}/z) \cdot \log \frac{p(\underset{\sim}{\theta}/z)}{p(\underset{\sim}{\theta})} \, d\underset{\sim}{\theta} \, dz$$

El problema planteado con $E^s(\underset{\sim}{n})$ por la existencia de cantidades de interés transformadas del primitivo vector paramétrico –problema ya tratado por Bernardo (1978) para experimentos simples– admite una definición en línea con la de la información esperada útil. Y así, siendo $\underset{\sim}{\psi} = \psi(\underset{\sim}{\theta})$ el vector de interés transformado del primitivo $\underset{\sim}{\theta}$, el investigador deberá cuantificar la información que el experimento $E^s(\underset{\sim}{n})$ proporciona sobre dicho nuevo vector. Surge así, en contraposición a la información esperada global anterior que podemos considerar como total, el concepto de información esperada global útil definido así:

DEFINICION. La información esperada global útil sobre $\underset{\sim}{\psi}$ proporcionada por el experimento $E^s(\underset{\sim}{n})$, cuando la densidad inicial es $p(\underset{\sim}{\theta})$ y la transformación es $\underset{\sim}{\psi} = \psi(\underset{\sim}{\theta})$, viene dada a través de la doble integral

$$I^{\psi}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} = \iint p(z) \cdot p(\underset{\sim}{\psi}/z) \cdot \log \frac{p(\underset{\sim}{\psi}/z)}{p(\underset{\sim}{\psi})} \, d\underset{\sim}{\psi} \, dz$$

donde $p(\underset{\sim}{\psi})$ es la densidad inicial deducida de $p(\underset{\sim}{\theta})$ por la transformación $\underset{\sim}{\psi} = \psi(\underset{\sim}{\theta})$; y $p(\underset{\sim}{\psi}/z)$ es la densidad posterior deducida de $p(\underset{\sim}{\theta}/z)$ por la transformación $\underset{\sim}{\psi} = \psi(\underset{\sim}{\theta})$.

El estudio de las propiedades satisfechas por las informaciones esperadas globales, útil y total, puede encontrarse en Veres (1981a). En resumen, ambas van a verificar ciertos criterios de aditividad que resultarán base para el desarrollo de un procedimiento que reduce el experimento global $E^s(\underset{\sim}{n})$ a un experimento simple, transformando las dos definiciones anteriores a la situación k-dimensional de la ya estudiada por Lindley (1956).

2. CALCULO DE LA INFORMACION ESPERADA GLOBAL

Es conocido que el cálculo práctico de la información esperada puede resultar de gran dificultad ante la complejidad del integrando que aparece en la definición (véase, por ejemplo, los resultados del modelo binomial en Bernardo (1976)). Si esta dificultad ya es notoria cuando la cantidad de interés es una variable aleatoria unidimensional y la muestra resultado del experimento $E(n)$ se extrae de un único espacio muestral, es evidente entonces la dificultad adicional que supone manejar un vector paramétrico como cantidad de interés y una muestra que es conjunto de ellas.

Precisamente para resolver este problema puede estudiarse el comportamiento de la información esperada en muestras grandes y así obtener fórmulas relativamente sencillas que la describan con más o menos grado de aproximitividad. En esta línea están los trabajos de Bernardo (1979) e Ibragimov & Has'Minsky (1973). Otro gran grupo de procedimientos se basa en estudiar la información esperada sobre cierta cantidad transformada mediante una biyección, apoyándose en la invarianza de la información esperada ante tales transformaciones.

Una de las mayores dificultades introducida por la existencia del experimento $E^S(\underset{\sim}{n})$ radica en que considera un vector tamaño muestral cuyas componentes no son necesariamente iguales. En caso contrario (con una afijación uniforme) la obtención de $I^{\underset{\sim}{\theta}}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\}$ podría simplificarse considerablemente suponiendo que z es una muestra de cierto tamaño k (indicado por la afijación) de la población L - dimensional $X_1 \, x \, X_2 \, x \ldots x \, X_L$. De esta forma, la determinación de la información esperada global quedaría encuadrada dentro de la teoría general, como un caso de L dimensiones. Sin embargo, esta situación no es la general, por lo que se hace necesario encontrar un procedimiento que permita reducir $E^S(\underset{\sim}{n})$ o establecer su equivalencia con ciertos experimentos simples. El objetivo de la presente sección consistirá en presentar dos de ellos.

I. OBTENCION DE LA INFORMACION ESPERADA GLOBAL POR EXTENSION DE LA CANTIDAD DE INTERES. Sea un vector tamaño muestral genérico $\underset{\sim}{n}'$ para el que definimos la matriz A siguiente:

$$A = \begin{pmatrix} \left.\begin{matrix} 1 & 0 & \ldots & 0 \\ & \cdots\cdots & & \\ 1 & 0 & \ldots & 0 \end{matrix}\right\} n_1 \\ \cdots\cdots \\ \cdots\cdots \\ \left.\begin{matrix} 0 & 0 & \ldots & 1 \\ & \cdots\cdots & & \\ 0 & 0 & \ldots & 1 \end{matrix}\right\} n_L \end{pmatrix}_{n \times L}$$

donde $n = n_1 + n_2 + \cdots + n_L$. Consideremos el nuevo vector paramétrico $\underset{\sim}{\mu} = A\ \underset{\sim}{\theta}$ que toma valores en el conjunto

$$\mu = \prod_{i=1}^{L} \Theta_i \times \Theta_i \times \overset{(n_i)}{\ldots\ldots} \times \Theta_i$$

Sea una transformación inversa cualquiera de la anterior, la definida, por ejemplo, por la matriz B de expresión:

$$B = \begin{pmatrix} 1 & 0 \overset{n_1}{\ldots\ldots} 0 & 0 & 0 \overset{n_2}{\ldots\ldots} 0 \ldots\ldots 0 & 0 \overset{n_L}{\ldots\ldots} 0 \\ 0 & 0 \ldots\ldots 0 & 1 & 0 \ldots\ldots 0 \ldots\ldots 0 & 0 \ldots\ldots 0 \\ & & \cdots\cdots & & \\ & & \cdots\cdots & & \\ 0 & 0 \ldots\ldots 0 & 0 & 0 \ldots\ldots 0 \ldots\ldots 1 & 0 \ldots\ldots 0 \end{pmatrix}_{L \times n}$$

verificándose que $\underset{\sim}{\theta} = B\ \underset{\sim}{\mu}$. Siendo $p(\underset{\sim}{\theta})$ la densidad de probabilidad que describe las opiniones iniciales que el investigador posee sobre $\underset{\sim}{\theta}$, las opiniones iniciales sobre el nuevo parámetro $\underset{\sim}{\mu}$ vendrán descritas por la correspondiente densidad transformada (supuestamente existente) $p(\underset{\sim}{\mu}) = A(p(\underset{\sim}{\theta}))$ que es degenerada al tener toda su masa contenida en un hiperplano de L dimensiones. Sea $\underset{\sim}{\mu}^*$ el subvector tal que $\underset{\sim}{\mu} = (\underset{\sim}{\theta}, \underset{\sim}{\mu}^*)$. Así pues, $\underset{\sim}{\mu}^*$ toma valores en el conjunto:

$$\mu^* = \prod_{i=1}^{L} \Theta_i \times \Theta_i \times \overset{n_i - 1}{\ldots\ldots} \times \Theta_i$$

Denotemos por $m(\underset{\sim}{\theta})$ a la densidad marginal

$$m(\underset{\sim}{\theta}) = \int p(\underset{\sim}{\mu})\, d\underset{\sim}{\mu}^*$$

LEMA 1

La densidad marginal $m(\underset{\sim}{\theta})$ deducida de $p(\underset{\sim}{\mu})$ coincide con $p(\underset{\sim}{\theta})$

Demostración.— La relación entre las funciones características de las variables $\underset{\sim}{\theta}$ y $\underset{\sim}{\mu}$ es:

$$\varphi_{\underset{\sim}{\theta},\, p(\underset{\sim}{\theta})}(t) = \varphi_{\underset{\sim}{\mu},\, p(\underset{\sim}{\mu})}(B't)$$

donde t es un vector columna $L \times 1$. Denotando por $E_{\omega,\, p(\omega)}$ a una esperanza definida sobre ω con la densidad de probabilidad $p(\omega)$, resulta:

$$\varphi_{\underset{\sim}{\theta},\, p(\underset{\sim}{\theta})}(t) = E_{\underset{\sim}{\theta},\, p(\underset{\sim}{\theta})}(e^{it'\underset{\sim}{\theta}}) = E_{\underset{\sim}{\mu},\, p(\underset{\sim}{\mu})}(e^{it'B\underset{\sim}{\mu}}) = E_{\underset{\sim}{\mu},\, p(\underset{\sim}{\mu})}(e^{it'\underset{\sim}{\theta}}) =$$

$$= E_{(\underset{\sim}{\theta},\, \underset{\sim}{\mu}^*),\, p(\underset{\sim}{\theta},\, \underset{\sim}{\mu}^*)}(e^{it'\underset{\sim}{\theta}}) = E_{\underset{\sim}{\theta},\, m(\underset{\sim}{\theta})}(e^{it'\underset{\sim}{\theta}}) = \varphi_{\underset{\sim}{\theta},\, m(\underset{\sim}{\theta})}(t) \quad \text{(c.q.d.)}$$

Apliquemos este resultado para encontrar un procedimiento que permita la obtención exacta de la información esperada global. Para ello veamos la equivalencia entre las siguientes situaciones: por una parte, realizar el experimento $E^s(\underset{\sim}{n})$ a fin de recabar información sobre el vector cantidad de interés $\underset{\sim}{\theta}$; por otra, realizar el experimento $E(E^s(\underset{\sim}{n})) = \{Z, \mu, p(z/\mu)\}$ consistente en la extracción de una unidad muestral de la población n-dimensional

$$Z = \prod_{i=1}^{L} X_i \times X_i \times \overset{n_i}{\dots} \times X_i$$

con objeto de obtener información sobre la cantidad de interés transformada $\underset{\sim}{\mu} = A\, \underset{\sim}{\theta}$. Notemos la existencia de las siguientes igualdades entre densidades:

$$p_{p(\underset{\sim}{\mu})}(z) = \iint p(z/\underset{\sim}{\theta})\, p(\underset{\sim}{\theta},\, \underset{\sim}{\mu}^*)\, d\underset{\sim}{\theta}\, d\underset{\sim}{\mu}^* = p_{p(\underset{\sim}{\theta})}(z)$$

$$p(z/\underset{\sim}{\mu}) = \prod_{i=1}^{L} \prod_{j=1}^{n_i} p\,(x_{ij}/\theta_i) = p(z/\underset{\sim}{\theta})$$

TEOREMA 1

$$I^{\underset{\sim}{\theta}}\, \{E^s(\underset{\sim}{n}),\, p(\underset{\sim}{\theta})\} = I^{\underset{\sim}{\mu}}\, \{E(E^s(\underset{\sim}{n}),\, p(\underset{\sim}{\mu})\}$$

Demostración.— De sus mismas definiciones:

$$I^{\underset{\sim}{\mu}}\{E(E^s(\underset{\sim}{n}), p(\underset{\sim}{\mu})\} = \iint p(\underset{\sim}{\mu}) \cdot p(z/\underset{\sim}{\mu}) \cdot \log \frac{p(z/\underset{\sim}{\mu})}{p(z)} \cdot d\underset{\sim}{\mu} \cdot dz =$$

$$= \iint p(\underset{\sim}{\mu}) \cdot p(z/\underset{\sim}{\theta}) \cdot \log \frac{p(z/\underset{\sim}{\theta})}{p(z)} \cdot d\underset{\sim}{\mu} \cdot dz = \text{(por Lema anterior)} =$$

$$= \iint p(\underset{\sim}{\theta}) \cdot p(z/\underset{\sim}{\theta}) \cdot \log \frac{p(z/\underset{\sim}{\theta})}{p(z)} \cdot d\underset{\sim}{\theta} \cdot dz = I^{\underset{\sim}{\theta}}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} \quad \text{(c.q.d.)}$$

Destaquemos que el Teorema 1 es aplicable indistintamente sobre informaciones esperadas globales, tanto totales como útiles. Por otra parte, y al ser la dimensión del parámetro transformado n, el resultado anterior reduce nuestro problema al caso n-dimensional del estudiado por Lindley (1956).

II. OBTENCION DE LA INFORMACION ESPERADA GLOBAL POR CONSIDERACION DE ESTADISTICOS SUFICIENTES EN TODOS LOS ESTRATOS. La situación expresada por la existencia de un experimento $E^s(\underset{\sim}{n})$ es también equivalente a la indicada por otro experimento realizado para obtener información también sobre $\underset{\sim}{\theta}$ y consistente en obtener una muestra unitaria de la población L-dimensional $Z = Z_1 \times Z_2 \times \ldots \times Z_L$, cuando Z_j es, a su vez, la población n_j-dimensional $Z_j = X_j \times X_j \times ..(n_j).. \times X_j$. Es evidente que el problema anterior quedará reducido a la situación L-dimensional estudiada por Lindley (1956) cuando podamos sustituir cada población Z_j por una población unidimensional.

Sean los experimentos $E^s(\underset{\sim}{n})$ y $E^s(\underset{\sim}{n} + 1)$ que se diferencian entre sí en que éste último obtiene una unidad adicional de muestreo. Sean $z(\underset{\sim}{n})$ y $z(\underset{\sim}{n} + 1)$ sus resultados respectivos. Podemos afirmar entonces que existirá un cierto estrato i_0-ésimo en el que el experimento componente $E(n_{i_0} + 1)$ del $E^s(\underset{\sim}{n} + 1)$ consistirá en la obtención de una unidad de muestreo adicional respecto al experimento $E(n_{i_0})$ componente i_0-ésimo de $E^s(\underset{\sim}{n})$. Así pues, existirá un cierto i_0 para el que $p(z(\underset{\sim}{n} + 1)/\underset{\sim}{\theta}) = p(x_{i_0 n_{i_0}} + 1/\theta_{i_0}) \cdot p(z(\underset{\sim}{n})/\underset{\sim}{\theta}$

Sea ahora el experimento $E(n_{i_0} + 1)^a/z(\underset{\sim}{n})$ consistente en obtener la unidad $(n_{i_0} + 1)$-ésima en el estrato X_{i_0}, supuestamente obtenidas y conocidas las demás muestras $z_i(n_i)$ $(i = 1, 2, \ldots, L)$. Tendremos, pues:

$$E(n_{i_0}+1)^a/z(n)=\{X_{i_0},\Theta_{i_0},p(x_{i_0\,n_{i_0}}/\theta_{i_0},z(\underset{\sim}{n}))\}$$

de forma que la información esperada que proporciona sobre $\underset{\sim}{\theta}$ será:

$$I^{\underset{\sim}{\theta}}\{E(n_{i_0}+1)^a/z(\underset{\sim}{n}),p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))\}=$$

$$=\iint p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))\cdot p(x_{i_0\,n_{i_0}+1}/\underset{\sim}{\theta},z(\underset{\sim}{n}))\cdot\log\frac{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}),x_{i_0\,n_{i_0}+1})}{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))}\cdot d\theta\cdot dx_{i_0 n_{i_0}+1}$$

El valor esperado sobre todas las posibles muestras $z(\underset{\sim}{n})$ de la información anterior recibirá el nombre de información esperada sobre $\underset{\sim}{\theta}$ proporcionada por $E(n_{i_0}+1)^a$ una vez que se ha realizado $E^s(\underset{\sim}{n})$. La denotaremos por

$$I^{\underset{\sim}{\theta}}\{E(n_{i_0}+1)^a/E^s(\underset{\sim}{n}),p(\underset{\sim}{\theta})\}$$

Con los anteriores experimentos puede demostrarse la existencia de una aditividad perfecta satisfecha por la información esperada global.

TEOREMA 2

Dados los experimentos $E^s(\underset{\sim}{n})$ y $E^s(\underset{\sim}{n}+1)$, una densidad inicial $p(\underset{\sim}{\theta})$ y suponiendo que la nueva unidad muestral que diferencia a ambos experimentos pertenece al estrato i-ésimo, se verifica la siguiente aditividad:

$$I^{\underset{\sim}{\theta}}\{E^s(\underset{\sim}{n}+1),p(\underset{\sim}{\theta})\}=I^{\underset{\sim}{\theta}}\{E^s(\underset{\sim}{n}),p(\underset{\sim}{\theta})\}+I^{\underset{\sim}{\theta}}\{E(n_{i_0}+1)^a/E^s(\underset{\sim}{n}),p(\underset{\sim}{\theta})\}$$

Demostración.– De sus definiciones respectivas resulta:

$$I^{\underset{\sim}{\theta}}\{E^s(\underset{\sim}{n}),p(\underset{\sim}{\theta})\}=\iint p(\underset{\sim}{\theta})\cdot p(z(\underset{\sim}{n}+1)/\underset{\sim}{\theta})\cdot\log\frac{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))}{p(\underset{\sim}{\theta})}\cdot d\underset{\sim}{\theta}\cdot dz(\underset{\sim}{n}+1)$$

$$I^{\underset{\sim}{\theta}}\{E(n_{i_0}+1)^a/E^s(\underset{\sim}{n}),p(\underset{\sim}{\theta})\}=\int p(z(\underset{\sim}{n}))\iint p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))\cdot$$

$$\cdot\,p\,(x_{i_0\,n_{i_0}+1}/\underset{\sim}{\theta},z(\underset{\sim}{n}))\cdot\log\frac{p(\underset{\sim}{\theta}/x_{i_0\,n_{i_0}+1},z(\underset{\sim}{n}))}{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))}\cdot d\underset{\sim}{\theta}\cdot dx_{i_0\,n_{i_0}+1}\cdot dz(\underset{\sim}{n})=$$

$$=\iint p(\underset{\sim}{\theta})\cdot p(z(\underset{\sim}{n}+1)/\underset{\sim}{\theta})\cdot\log\frac{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}+1))}{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))}\cdot d\underset{\sim}{\theta}\cdot dz\,(\underset{\sim}{n}+1)$$

Sumando ambas informaciones resulta finalmente:

$$\iint p(\underset{\sim}{\theta}) \cdot p(z(\underset{\sim}{n}+1)/\underset{\sim}{\theta}) \cdot \log \frac{p(\underset{\sim}{\theta}/z(\underset{\sim}{n}+1))}{p(\underset{\sim}{\theta})} \cdot d\underset{\sim}{\theta} \cdot dz(\underset{\sim}{n}+1) =$$

$$= I^{\underset{\sim}{\theta}} \{E^s(\underset{\sim}{n}+1), p(\underset{\sim}{\theta})\} \quad \text{(c.q.d.)}$$

El resultado en verdad interesante a nuestros propósitos es el contenido en el siguiente:

COROLARIO 1

Siendo $z(\underset{\sim}{n})$ suficiente para $z(\underset{\sim}{n}+1)$ con respecto a $\underset{\sim}{\theta}$ en todos los estratos, esto es, verificándose que $p(\underset{\sim}{\theta}/z(\underset{\sim}{n}), x_{in_i+1}) = p(\underset{\sim}{\theta}/z(\underset{\sim}{n}))$ para todo $i = 1, ...,$ entonces:

$$I^{\underset{\sim}{\theta}} \{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} = I^{\underset{\sim}{\theta}} \{E^s(\underset{\sim}{n}+1), p(\underset{\sim}{\theta})\}$$

Demostración.— Verificándose la hipótesis de igualdad entre dichas densidades, $I^{\underset{\sim}{\theta}} \{E(n_i+1)^a/E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\}$ es nula en todos los estratos (Veres, 1981a). Así pues, e independientemente del estrato en el que se extraiga la última unidad muestral del experimento $E^s(\underset{\sim}{n}+1)$, se sigue el resultado del Teorema 2. (c.q.d.)

Este corolario establece una importante consecuencia: no hay pérdida en la información esperada global cuando la investigación se ciñe en cada estrato a la observación de un estadístico suficiente respecto todo el vector de interés $\underset{\sim}{\theta}$.

Siguiendo idénticas técnicas de demostración que las utilizadas para demostrar el Teorema 2 y el consiguiente Teorema 1, pueden demostrarse idénticos resultados para la información esperada global útil, lo que permite aplicar a ambas informaciones el procedimiento que ahora va a desarrollarse.

Denotemos por $E(\underset{\sim}{1}, S)$ al experimento consistente en la obtención de una unidad muestral de un espacio L - dimensional S en el que toma valores un conjunto de L estadísticos suficientes respecto todo el vector aleatorio $\underset{\sim}{\theta}$. El Corolario 2 establece la equivalencia entre este experimento y el $E^s(\underset{\sim}{n})$. Asi pues, sustituyendo los espacios muestrales Z_i

por sendos S_i en los que toman valores un conjunto de L estadísticos φ_i tales que:

$$p(\underset{\sim}{\theta}/z_1, ..., z_{i-1}, \varphi_i, z_{i+1}, ..., z_L) = p(\underset{\sim}{\theta}/z) \quad \text{para todo } i = 1, 2, ..., L$$

resultará ser:

$$I^{\theta}_{\sim}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} = I^{\theta}_{\sim}\{E(\underset{\sim}{1}, S), p(\underset{\sim}{\theta})\}$$

$$I^{\psi}_{\sim}\{E^s(\underset{\sim}{n}), p(\underset{\sim}{\theta})\} = I^{\psi}_{\sim}\{E(\underset{\sim}{1}, S), p(\underset{\sim}{\theta})\}$$

De esta forma la obtención de la información esperada global quedará reducida a la situación L-dimensional estudiada por Lindley (1956)

3. EL MODELO NORMAL

La extensión del modelo normal a un experimento de la forma $E^s(\underset{\sim}{n})$ supone considerar un conjunto de L subpoblaciones sobre las que se definen sendas densidades normales $N(\theta_i, \sigma_i)$ $(i = 1, 2, ..., L)$, siendo conocidas todas las σ_i. La densidad inicial que describe el comportamiento del vector de interés $\underset{\sim}{\theta}$ es también multinormal. Así pues, estableceremos como hipótesis básicas para este modelo las dos siguientes:

1ª $p(x_i/\theta_i)$ es $N(\theta_i, \sigma_i)$, $i = 1, 2, ..., L$

2ª $p(\underset{\sim}{\theta})$ es multinormal de vector media $\underset{\sim}{\theta}_0$ y matriz H_0 de elementos σ^0_{ij}. De esta forma el experimento $E^s(\underset{\sim}{n})$ se definirá a partir de la familia de L experimentos $E(n_i)$, donde:

$$E_i = \{X_i, \Theta_i, N(\theta_i, \sigma_i)\} \quad i = 1, 2, ..., L$$

I. *Obtención de la información esperada por extensión de la cantidad de interés.* La hipótesis 1ª permite considerar a la muestra global z como una muestra unitaria de una población normal n-dimensional, de media $A \cdot \underset{\sim}{\theta}$ y matriz diagonal de momentos M, donde:

$$M = \begin{pmatrix} \sigma_1^2 & 0 \overset{n_1}{\ldots\ldots} 0 & \ldots & 0 & 0 \overset{n_L}{\ldots\ldots} 0 \\ & & \ldots\ldots & & \\ 0 & 0 \ldots & \sigma_1^2 \ldots & 0 & 0 \ldots 0 \\ & & \ldots\ldots & & \\ & & \ldots\ldots & & \\ 0 & 0 \ldots 0 & \ldots \sigma_L^2 & & 0 \ldots 0 \\ & & \ldots\ldots & & \\ 0 & 0 \ldots 0 & \ldots 0 & & 0 \ldots \sigma_L^2 \end{pmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} n_1 \\ \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} n_L \end{array}$$

Sea $\underset{\sim}{\mu} = A \cdot \underset{\sim}{\theta}$ el nuevo parámetro extendido. Entonces, la distribución de $\underset{\sim}{\mu}$ será una normal n-dimensional degenerada, de vector media $A \,.\, \underset{\sim}{\theta}_0$ y matriz de momentos singular AH_0A'. Bajo la transformación anterior y aplicando el resultado de Lindley (1956) y el Teorema 1, se obtiene finalmente:

$$I^{\underset{\sim}{\theta}} \{E^s(\underset{\sim}{n}), N^L(\underset{\sim}{\theta}_0, H_0)\} = I^{\underset{\sim}{\mu}} \{E(E^s(\underset{\sim}{n}), N^n(A \cdot \underset{\sim}{\theta}_0, AH_0A')\} =$$

$$= \frac{1}{2} \log \frac{|AH_0A' + M|}{|M|} \tag{1}$$

II. *Obtencioń de la información esperada por consideración de estadísticos suficientes en todos los estratos.* Sea N una matriz $L \times L$ diagonal cuyos elementos no nulos son los distintos tamaños de las muestras extraídas en sendos estratos. Denotemos por H a otra matriz diagonal $L \times L$ cuyos elementos no nulos son las varianzas σ_i^2, siendo $R = H^{-1}$. Denotemos por T a la matriz de precisión de la densidad inicial $p(\underset{\sim}{\theta})$, esto es, $T = H_0^{-1}$. Sea $\underset{\sim}{\bar{x}}$ el vector columna $L \times 1$ cuyas componentes son las medias muestrales de las submuestras extraídas en los estratos respectivos. Precisamente, puede considerarse al vector $\underset{\sim}{\bar{x}}$ como el conjunto de L estadísticos suficientes respecto todo el vector de interés. Es conocido (DeGroot, 1970) que el vector $\underset{\sim}{\bar{x}}$ se distribuye según una multinormal de media $\underset{\sim}{\theta}$ y matriz de precisión NR.

Aplicando el segundo de los procedimientos de obtención para la información esperada global, ésta podrá obtenerse estudiando la información contenida en una muestra unitaria del espacio $\underset{\sim}{\bar{X}} = \bar{X}_1 \times \bar{X}_2 \times \ldots \times \bar{X}_L$ en el que toma valores el conjunto de estadísticos suficientes $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_L$. Así pues, por el Corolario 1 y teniendo en cuenta la expresión de la información esperada sobre $\underset{\sim}{\theta}$ proporcionada por una unidad muestral en el modelo normal n-dimensional (Lindley, 1956), resulta:

$$I^{\theta}\{E^s(\underset{\sim}{n}), N^L(\underset{\sim}{\theta}_0, H_0)\} = I^{\theta}\{E^s(\underset{\sim}{1}, \bar{X}), N^L(\underset{\sim}{\theta}_0, H_0)\} =$$

$$= \frac{1}{2} \log \frac{|H_0 + (NR)^{-1}|}{|(NR)^{-1}|} = \frac{1}{2} \log \frac{|H_0 N + H|}{|H|} \qquad (2)$$

La igualdad de las expresiones (1) y (2) puede verse en Veres (1981b).

III. *Obtención de las informaciones esperadas globales sobre θ y θ_i.* Apliquemos ahora el segundo de los procedimientos para la obtención de una información esperada global útil. Supongamos, pues, que bajo las hipótesis del modelo normal se realiza el experimento $E^s(\underset{\sim}{n})$ a fin de adquirir información sobre la cantidad aleatoria unidimensional transformada $\theta = J'\underset{\sim}{\theta}$ con $J' = (h_1, h_2, ..., h_L)$. Consideremos nuevamente el vector $\bar{\underset{\sim}{x}}$ que se distribuye según una multinormal $N^L(\underset{\sim}{\theta}, HN^{-1})$. Así pues, $p(\underset{\sim}{\theta}/\bar{\underset{\sim}{x}})$ será otra multinormal $N^L((H_0^{-1} + NH^{-1})^{-1}(H_0^{-1}\underset{\sim}{\theta}_0 + NH^{-1} \cdot \bar{\underset{\sim}{x}}), (H_0^{-1} + NH^{-1})^{-1})$. De acuerdo con el segundo de los procedimientos para la obtención de la información esperada global, resulta ser: $I^{\theta}\{E^s(\underset{\sim}{n}), N^L(\underset{\sim}{\theta}_0, H_0)\} = I^{\theta}\{E(\underset{\sim}{1}, \bar{X}), N^L(\underset{\sim}{\theta}_0, H_0)\}$. Por otra parte, siendo θ una combinación lineal de las componentes del vector $\underset{\sim}{\theta}$, deducimos que $p(\theta) = N(J'\underset{\sim}{\theta}_0, J'H_0J)$ y que $p(\theta/\bar{\underset{\sim}{x}}) = N(J'(H_0^{-1} + NH^{-1})^{-1}(H_0^{-1}\underset{\sim}{\theta}_0 + NH^{-1}, J'(H_0^{-1} + NH^{-1})^{-1}J)$. Demostremos los siguientes lemas:

LEMA 2

Bajo la notación desarrollada, $p(\bar{\underset{\sim}{x}}) = N^L(\underset{\sim}{\theta}_0, H_0 + HN^{-1})$

Demostración.– En efecto,

$$p(\bar{\underset{\sim}{x}}) = \int p(\bar{\underset{\sim}{x}}/\underset{\sim}{\theta}) \cdot p(\underset{\sim}{\theta}) \cdot d\underset{\sim}{\theta} = \int N^L(\underset{\sim}{\theta}, HN^{-1}) \cdot N^L(\underset{\sim}{\theta}_0, H_0) \cdot d\underset{\sim}{\theta}$$

Denotemos por $P = (HN^{-1})^{-1}$ y $P_0 = H_0^{-1}$ a las respectivas matrices de precisión de las densidades $p(\bar{\underset{\sim}{x}}/\underset{\sim}{\theta})$ y $p(\underset{\sim}{\theta})$. Resultará:

$$p(\bar{\underset{\sim}{x}}) = \frac{\sqrt{|PP_0|}}{(2\pi)^L} \int e^{-\frac{1}{2}(\bar{\underset{\sim}{x}} - \underset{\sim}{\theta})' P(\bar{\underset{\sim}{x}} - \underset{\sim}{\theta}) - \frac{1}{2}(\underset{\sim}{\theta} - \underset{\sim}{\theta}_0)' P_0(\underset{\sim}{\theta} - \underset{\sim}{\theta}_0)} \cdot d\underset{\sim}{\theta}$$

El exponente puede ponerse como suma de $A + B$, siendo:

$$A = \{\underset{\sim}{\theta} - (P + P_0)^{-1}(P_0\underset{\sim}{\theta}_0 + P\bar{\underset{\sim}{x}})\}'(P + P_0)\{\underset{\sim}{\theta} - (P + P_0)^{-1}(P_0\underset{\sim}{\theta}_0 + P\bar{\underset{\sim}{x}})\} =$$

$$= \{P_0(\underset{\sim}{\theta} - \underset{\sim}{\theta}_0) + P(\underset{\sim}{\theta} - \underset{\sim}{\bar{x}})\}' (P + P_0)^{-1} \{P_0(\underset{\sim}{\theta} - \underset{\sim}{\theta}_0) + P(\underset{\sim}{\theta} - \underset{\sim}{\bar{x}})\}$$

$$B = (\underset{\sim}{\bar{x}} - \underset{\sim}{\theta}_0)' P(P + P_0)^{-1} P_0 (\underset{\sim}{\bar{x}} - \underset{\sim}{\theta}_0)\}$$

Así pues,

$$p(\underset{\sim}{\bar{x}}) = \frac{\sqrt{|PP_0|}}{(2\pi)^L} \cdot e^{-½B} \int e^{-½A} \cdot d\underset{\sim}{\theta} = \frac{\sqrt{|PP_0|}}{(2\pi)^L} \cdot e^{-½B} \cdot \frac{(\sqrt{2\pi})^L}{\sqrt{|P + P_0|}} =$$

$$= \frac{1}{(\sqrt{2\pi})^L} \sqrt{\frac{|PP_0|}{|P + P_0|}} \cdot \exp(-½ (\underset{\sim}{\bar{x}} - \underset{\sim}{\theta}_0)' P(P + P_0)^{-1} P_0 (\underset{\sim}{\bar{x}} - \underset{\sim}{\theta}_0))$$

esto es,

$$p(\underset{\sim}{\bar{x}}) = N^L (\underset{\sim}{\theta}_0, (P(P + P_0)^{-1} P_0)^{-1}) = N^L (\underset{\sim}{\theta}_0, H_0 + NH^{-1})$$

(c.q.d.)

LEMA 3

$$\int p(\underset{\sim}{\bar{x}}) \cdot p(\theta/\underset{\sim}{\bar{x}}) \cdot d\underset{\sim}{\bar{x}} = N^L (J' \underset{\sim}{\theta}_0, J'H_0 J)$$

Demostración.– La función característica de la distribución de θ definida a partir de $p(\underset{\sim}{\bar{x}})$ y $p(\theta/\underset{\sim}{\bar{x}})$ viene dada por

$$\varphi(t) = \int e^{it} \left\{ \int p(\underset{\sim}{\bar{x}}) \cdot p(\theta/\underset{\sim}{\bar{x}}) \cdot d\underset{\sim}{\bar{x}} \right\} d\theta =$$

$$= \int p(\underset{\sim}{\bar{x}}) \left\{ \int e^{it\theta} N(J'(H_0^{-1} + NH^{-1})^{-1} (H_0^{-1} \underset{\sim}{\theta}_0 + NH^{-1} \cdot \underset{\sim}{\bar{x}}),\ J' (H_0^{-1} +\right.$$

$$\left. + NH^{-1})^{-1} J) \cdot d\theta \right\} d\underset{\sim}{\bar{x}} =$$

$$= e^{itJ'(H_0^{-1} + NH^{-1})^{-1} H_0^{-1} \underset{\sim}{\theta}_0 - ½ t^2 J'(H_0^{-1} + NH^{-1})^{-1} J} \cdot$$

$$\int e^{itJ'(H_0^{-1} + NH^{-1})^{-1} NH^{-1} \cdot \underset{\sim}{\bar{x}}} \cdot p(\underset{\sim}{\bar{x}}) \cdot d\underset{\sim}{\bar{x}} = \text{(por Lema anterior)} =$$

$$= \exp(itJ'\underset{\sim}{\theta}_0 - ½\, t^2 J'(I + H_0 NH^{-1}) (H_0^{-1} + NH^{-1})^{-1} J) =$$

$$= \exp(itJ'\underset{\sim}{\theta}_0 - ½\, t^2 J'H_0 J) \quad \text{(c.q.d.)}$$

A partir de los lemas anteriores puede obtenerse la información esperada global sobre θ proporcionada por el experimento $E^s(\underset{\sim}{n})$ y para el modelo normal.

TEOREMA 3

En el modelo normal, siendo $\theta = J'\underset{\sim}{\theta}$

$$I^{\theta}\{E^{s}(\underset{\sim}{n}), N^{L}(\underset{\sim}{\theta}_0, H_0)\} = ½ \log \frac{J'H_0J}{J'(H_0^{-1} + NH^{-1})^{-1}J}$$

Demostración.– El resultado se sigue de la siguiente cadena de igualdades:

$$I^{\theta}\{E^{s}(\underset{\sim}{n}), N^{L}(\underset{\sim}{\theta}_0, H_0)\} = I^{\theta}\{E(\underset{\sim}{1}, \bar{X}), N^{L}(\underset{\sim}{\theta}_0, H_0)\} =$$

$$= \log \sqrt{\frac{J'H_0J}{J'(H_0^{-1} + NH^{-1})^{-1}J}} - \frac{1}{2}\iint p(\underset{\sim}{\bar{x}}) \cdot p(\theta/\underset{\sim}{\bar{x}}) \cdot$$

$$\cdot \frac{(\theta - E_{p(\theta/\underset{\sim}{\bar{x}})}(\theta))^2}{V_{p(\theta/\underset{\sim}{\bar{x}})}(\theta)} \cdot d\theta \cdot d\underset{\sim}{\bar{x}} + \frac{1}{2}\iint p(\underset{\sim}{\bar{x}}) \cdot p(\theta/\underset{\sim}{\bar{x}}) \cdot$$

$$\cdot \frac{(\theta - E_{p(\theta)}(\theta))^2}{V_{p(\theta)}(\theta)} \cdot d\theta \cdot d\underset{\sim}{\bar{x}} = (\text{por Lema } 3) = \frac{1}{2} \log \frac{J'H_0J}{J'(H_0^{-1} + NH^{-1})^{-1}J}$$

y donde $E_{p(\omega)}$ y $V_{p(\omega)}$ denotan, respectivamente, la media y la varianza calculadas sobre la distribución $p(\omega)$ de la variable aleatoria ω.(c.q.d.)

Este resultado permite cuantificar la información que todo el experimento global $E^{s}(\underset{\sim}{n})$ proporciona sobre una determinada componente del vector $\underset{\sim}{\theta}$ considerada como nueva cantidad de interés. Así, sea el vector J definido como: $J' = (0 \ldots 0\ 1\ 0 \ldots 0)$, siendo $\theta_i = J'\underset{\sim}{\theta}$. Notemos que $J'H_0J = \sigma_{ii}^0$ y que $J'(H_0^{-1} + NH^{-1})^{-1}J = a_{ii}$. Siendo m_{ii} el elemento (i, i) de la matriz $H_0^{-1} + NH^{-1}$, que es la matriz de precisión de la densidad $p(\underset{\sim}{\theta}/\underset{\sim}{\bar{x}})$, resultará ser:

$$J'(H_0^{-1} + NH^{-1})^{-1}J = \frac{|\operatorname{Adj} m_{ii}|}{|H_0^{-1} + NH^{-1}|} = \frac{\left|\operatorname{Adj}\left(\frac{\operatorname{Adj} \sigma_{ii}^0}{H_0} + \frac{n_i}{\sigma_i^2}\right)\right|}{|H_0^{-1} + NH^{-1}|}$$

De ahí que por el Teorema 3 resulte evidente el siguiente:

COROLARIO 2

Siendo m_{ii} el elemento (i, i) de la matriz $H_0^{-1} + NH^{-1}$ y para el modelo normal:

$$I^{\theta_i}\{E^s(\underset{\sim}{n}), N^L(\underset{\sim}{\theta}_0, H_0)\} = \frac{1}{2}\log\frac{\sigma_{ii}^0 |H_0^{-1} + NH^{-1}|}{|\text{Adj } m_{ii}|}$$

REFERENCIAS

BERNARDO, J.M. (1976). *The use of information in the design and analysis of scientific experimentation.* Ph. D. Tesis. Universidad de Londres.

BERNARDO, J.M. (1978). Una medida de la información útil proporcionada por un experimento. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales de Madrid,* 72. 419-440.

BERNARDO, J.M. (1979). Comportamiento asintótico de la información proporcionada por un experimento. *Revista de la Real Academia de Ciencias Exactas Físicas y Naturales de Madrid,* 73. 491-502.

DeGROOT, M.H. (1970). *Optimal Statistical Decisions.* New York: McGraw.

IBRAGIMOV, I.A. & HAS'MINSKY, R.Z. (1973). On the information in a sample about a parameter. *2nd. Internat. symp. Information Theory.* (Petrov & Csaki eds.). Budapest: Akadémiaikiadó.

LINDLEY, D.V. (1956). On a Measure of the Information provided by an Experiment. *Ann. Math. Statisi.* 27, 986-1005.

SHANNON, C.E. (1948). A Mathematical Theory of Communication. *Bell System Tech. J.* 27, 379-423, 623-56.

VERES, E. (1983). Información esperada proporcionada por un experimento asociado a un diseño estratificado. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales de Madrid,* 77. 569-595.

VERES, E. (1981b). *Diseño Bayesiano de muestras, posiblemente estratificadas, cuando la utilidad terminal es función de la información conseguida.* Tesis, Universidad de Valencia.